

English Longitudinal Study of Ageing (ELSA) User Guide for the Wave 2 (2004-5) Core Dataset Version 2

This document accompanies the second version of the interview data from Wave 2 (2004-5) of the English Longitudinal Study of Ageing (ELSA). The guide provides:

- An outline of the sampling and methodology,
- Content of the ELSA interview,
- A description of the ELSA Wave 2 core dataset (including key variables)
- Documentation of the ELSA Wave 2 main interview CAPI questionnaire
- ELSA Wave 2 core self completion questionnaire
- Code book and edit instructions
- Field documents (including cognitive function booklet, advance letters sent to respondents etc.)
- Interviewer instructions

The following data and documentation are deposited at the UK Data Archive (www.esds.ac.uk).

Table 1 ELSA Wave 2 Files deposited at the UK Data Archive

Title	Format
Wave 2 Core Data v2	Various statistical packages
Wave 2 Nurse Data v2	Various statistical packages
Wave 2 Financial Derived Variables (DVs) Data	Various statistical packages
Wave 2 Mortgage Data	Various statistical packages
Wave 2 Ryff Self Completion Data	Various statistical packages
Wave 2 Pension Grid	Various statistical packages
Wave 2 Core Data User Guides and Documentation	Word (available as part of PDF documentation)
Wave 2 Financial DVs Variable Relationships	Excel
Wave 2 Pension Grid Corresponding Variables	Excel
<i>Retirement, health and relationships of the older population in England: The 2004 English Longitudinal Study of Ageing. Report on the Wave 2 data.</i>	Link to website: www.ifs.org.uk/elsa

The following data and documentation are also available from the UK Data Archive:

Table 2 Other ELSA Files deposited at the UK Data Archive

Title	Format
ELSA Wave 1 Core Data	Various statistical packages
ELSA Wave 1 Financial Derived Variables (DVs) Data	Various statistical packages
ELSA Wave 1 Pension Wealth DVs Data	Various statistical packages

ELSA Wave 1: Core Data User Guides and Documentation	Word (available as part of PDF documentation)
ELSA Wave 1 Variable Relationships for Financial DVs	Excel
ELSA Wave 1 Financial DVs Variable Name Lookup	Text file (details variable name truncation that may occur with SPSS data)
ELSA Wave 1 Pension Wealth Variable Name Lookup	Text file (details variable name truncation that may occur with SPSS data)
<i>Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing. Report on the Wave 1 data.</i>	Link to website: www.ifs.org.uk/elsa
ELSA Wave 0 (i.e. Health Survey for England, HSE) data for respondents eligible for ELSA (<i>There is a dataset for each year of HSE that was sampled for ELSA Wave 1, i.e. 1998, 1999, 2001 - see Sample Design section below</i>)	Various statistical packages
ELSA Wave 0 (HSE) Common Variables Data (<i>Contains all respondents from the three HSE years sampled; only contains those variables included in all three survey years</i>)	Various statistical packages
ELSA Wave 0 Data User Guide	Word (available as part of PDF documentation)
ELSA Index File (containing summary data and all serial numbers for all eligible respondents from each wave)	Various statistical packages
ELSA Index File User Guide	Word (available as part of PDF documentation)

BACKGROUND AND AIMS

The English Longitudinal Study of Ageing (ELSA) is a study of people aged 50 and over and their younger partners, living in private households in England. The sample was drawn from households that had previously responded to the Health Survey for England (HSE) in 1998, 1999 or 2001.

Every two years we hope to interview the same group of people to measure change in their health, economic and social circumstances. ELSA can complete the picture of what it means to grow older in the new century, and help us understand what accounts for the variety of patterns that are seen.

This User Guide relates to data deposited for the interviews from the second Wave of ELSA, which were carried out between June 2004 and July 2005. This data can be analysed cross-sectionally, or longitudinally in conjunction with ELSA Wave 1 data and HSE Wave 0 data (both of which are in the data archive).

The ELSA Wave 2 interview covered a wide range of topics. It was similar to the questionnaire used in Wave 1, although every module was reviewed to ensure that it would provide data that measured change over time. This was achieved by repeating

some measures exactly (for example, to measure income and assets), by asking directly about change (for example, to capture perceived changes in memory and concentration) and by adapting questions to allow people to update or amend past responses (for example, about work, pensions and specific health conditions).

The Wave 2 interview was also expanded to answer a variety of additional research questions. The new items included: quality of health care received; household spending on leisure, clothing and transfers; perceptions of deprivation relative to others; perceptions of ageing; levels of literacy; perceived effort and reward for care-giving; and voluntary activities.

Core sample members who completed a main interview were also offered a nurse visit. This was similar to the one that many respondents had as part of the Health Survey for England in 1998, 1999 or 2001 and included tests of blood pressure, lung function, blood tests, anthropometric measures and physical performance measures. This data (and accompanying documentation) is also available from the archive.

ELSA is the result of collaboration between University College London, the Institute of Fiscal Studies (IFS), and the National Centre for Social Research (NatCen). Other academic collaborators based at the Universities of Cambridge, Exeter and East Anglia provided expert advice on specific modules.

Funding for the first four waves of ELSA has been provided by the US National Institute on Aging, and a consortium of British Government departments, specifically: Department for Education and Skills, Department of Environment, Food and Rural Affairs, Department of Work and Pensions, HM Treasury, HMRC (formerly Inland Revenue), Office of the Deputy Prime Minister and Office for National Statistics. Many of the measures adopted in ELSA are comparable with measures used in the US Health Retirement Study (HRS) and the Survey of Health and Retirement in Europe (SHARE).

ETHICAL CLEARANCE

Ethical approval for ELSA Wave 2 was granted from the Multicentre Research and Ethics Committee (MREC).

DEVELOPMENT AND PILOTING

Two pilots of the computer-assisted personal interview (CAPI) instrument, the self-completion questionnaire and associated documents were conducted in August 2003 and January 2004. All pilot respondents were drawn from households which participated in the ELSA Wave 1 pilots [KC to find out]. The pilots tested the fieldwork procedure and interview content. An approach to dependent interviewing was developed – that is, feeding information from a past interview into the current one.

SAMPLE DESIGN

The ELSA sample has been designed to represent people aged 50 and over, living in private households in England. Three years of the Health Survey for England (HSE) were selected as the sampling frame: 1998, 1999 and 2001. These years were chosen because they were recent and could provide a sufficiently large sample size. ELSA used the core samples for these years, all of which were nationally representative. The HSE 1999 sample design also included a boost sample that represented ethnic minorities. Because of funding constraints, it was not possible to follow-up the boost sample and it was discarded. Together these three HSE years contained 23,132 responding households.

Households were removed from the HSE sampling frame for ELSA Wave 1 if it was known that there was no adult of 50 years or older in the household who had agreed to be re-contacted at some time in the future. Individuals in the remaining households provided the basis for the ELSA Wave 1 sample (11,578 households containing 18,813 eligible individuals). The Wave 1 Technical Report provides more details.

Sample Types

Within households there were four different types of individual who were eligible to be invited to take part in the study: core sample members, core partners, younger partners and new partners. The variable **FINSTAT** determines the sample type of the respondent. Each type of respondent is described in Box 1.

The Wave 1 interview took place in 2002–2003 and provided the baseline for the study. Eligible sample members who responded at this stage were renamed '**core members**' to distinguish them as the core element of the continuing ELSA sample. They were eligible for the main interview in Wave 2 unless they had since died, had explicitly asked not to be re-contacted at the end of their first ELSA interview, or had moved out of Britain. Therefore, unlike Wave 1, respondents were eligible if they had moved to Wales or Scotland. Core members who completed a main interview in Wave 2 were also eligible for a nurse visit.

Several other categories of individuals were also eligible for an interview (but not a nurse visit) at Wave 2. These were the partners of core members (core partners, new partners or young partners, as described in Box 1).

Box 1 Summary of the eligibility criteria for the Wave 2 ELSA interview

Core members (C1CM) were individuals who had been living within the household at the time of the HSE interview in 1998, 1999 or 2001, were born on or before 29 February 1952 and were subsequently interviewed as part of Wave 1 at a private residential address in England. They were not eligible if they had since died, asked not to be revisited or moved out of Britain.

Core partners (C1CP) were individuals who, like core members, had been living within the household at the time of the HSE interview in 1998, 1999 or 2001 and were born on or before 29 February 1952. However they were *not* interviewed as part of Wave 1, so missing the baseline survey. Consequently, they were only approached by virtue of their being the partner of a core member.

Young partners (C1YP) were the cohabiting spouses or partners of eligible sample members, who were living within the household at the time of the HSE, and were still cohabiting with the sample member at the time of the Wave 1 interview. They were born *after* 29 February 1952. Most, but not all, young partners took part in a Wave 1 interview.

New partners (C1NP1 if identified in Wave 1; C1NP2 if identified in Wave 2) were the cohabiting spouses or partners of eligible sample members at the time of *either the first or second* ELSA interview, who had joined the household since the original HSE interview.

Core partners, young partners and new partners who had been identified in Wave 1 were eligible for a full Wave 2 interview even if they were no longer living with a core member at the time of the second interview. That is to say, we attempted to interview all partners who had been living with a core member at the time of an ELSA interview and had been separated or divorced from them, or had been widowed, so that we could understand their circumstances after this event had occurred. The only circumstances in which partners who had separated from the core member was not

approached was if they had died, had explicitly asked not to be re-contacted at the end of their first ELSA interview, had left Britain or had moved into an institution. Ex-partners of core members will only be followed up once after leaving the core member's household.

In version 1 (2?) of the Wave 2 deposited dataset there are 8780 core members, 57 core partners, 501 younger partners, and 94 new partners.

At Wave 2, two further types of interview were conducted with specific sub-populations. An '**end of life interview**' (also referred to as Exit interview) was sought with a relative, friend or carer of any core members who had died since responding to the first ELSA interview. An '**institution interview**' was sought with core members who had moved from a private household at the first ELSA interview into a residential care home or similar institution, or with a proxy who could respond on their behalf. The data collected during these two types of interviews is not included in the ELSA Wave 2 Core dataset, but may be deposited in the archive at a later date.

FIELDWORK

Fieldwork for the second wave of ELSA began in June 2004 and spanned 14 months, finishing in July 2005. Each eligible individual within a household was sent an advance letter inviting them to take part. Interviewers then visited the households to explain the study and to interview willing individuals straight away, or to make appointments to call at a convenient time. A number of different approaches were used to encourage participation among the sample (see the ELSA Wave 1 report).

ACHIEVED SAMPLE

The ELSA Wave 2 fieldwork produced 9432 productive interviews. As explained above, 8780 of these were with core sample members. Of all core sample member interviews, 38 people only completed a partial interview in person and 92 were interviewed by proxy (see variable **W2INDOUT**). In addition 57 productive interviews were conducted with core partners, 501 with younger partners, and 94 with new partners.

The interviewer made contact with 97% of the households that were issued for ELSA Wave 2 (the household contact rate). The largest component (77%) of non-response was a result of refusals. Though many people who had moved were traced from their Wave 1 residence, 11% of non-responders were individuals who could not be found. This is slightly higher than Wave 1, where those who had moved and could not be traced constituted 10% of issued Wave 1 non-respondents.

82% of those who completed a Wave 1 interview and were eligible for a Wave 2 interview as an ELSA 'core member' took part in the survey.

More information about the response rates is provided in the Wave 2 Technical Report and the Wave 2 report of the findings (see www.ifs.org.uk/elsa).

Response to key sections

In addition to the overall level of response, an analysis of the response to key sections (or modules) of the survey questionnaire was conducted. Not all modules required responses at an individual level. The household demographics and housing modules were asked at the household level, while the income and assets module was asked at the financial-unit level (see section on Survey Content). Table 3 shows the responses at the appropriate level for the three key modules of the main questionnaire, and for the nurse visit conducted in Wave 2 after the main interview.

Table 3 **Response rates to key modules**

Section	Total eligible	Level	Response rate %
Housing	6,246	Household	99.9
Income and assets	6,712	Financial unit	99.0
Self-completion	9,307	Individual	89.8
Nurse visit	8,688	Individual	88.2

The response rate for the housing, income and assets modules was very high and similar to the rates achieved in Wave 1. Response rates for the self-completion module (again similar to Wave 1) and nurse visit were good in survey terms. Further information about weighting to address non-response to the nurse visit, to the self-completion module and the blood sample collection is given the Wave 2 report (see www.ifs.org.uk/elsa). In addition, non-response to specific items in the interview, including economic variables, was very low, as it had been in Wave 1.

WEIGHTING

The variable in the dataset to be used for weighting is **W2WGT**. Weights were calculated for core sample members only (including proxy and partial interviews), as this is the sample of interest. All other non-sample individuals that were interviewed (i.e. core, new and younger partners) have a weight of zero. When running weighted analyses, researchers should remember to exclude non-sample members in the unweighted base, if quoted. The data for partners can be used as characteristics of the core members (i.e. supplementary information).¹

Where possible we recommend that analysis be conducted on weighted data since this should help to reduce bias from non-random non-response. The aim of weighting is to take account of any bias from non-response. That is, non-response weighting corrects for systematic differences in response rates across sub-groups. The equal probability design of the HSE samples, and the fact that the ELSA sample included all eligible adults from the HSE, eliminates any need for weights to account for selection probabilities. However, non-response at HSE, refusals to be re-interviewed post-HSE and non-response at ELSA Waves 1 and 2 all have the potential to make the ELSA respondent sample unrepresentative of those groups less likely to respond.

In addition, the original complex sample design of the HSE samples has to be considered. The HSE sample design is described in the ELSA Wave 1 Technical Report.

An analysis of response at Wave 2 focussed on the 10,770 core members who responded at Wave 1 and were *eligible* for interview at Wave 2. (Ineligible persons included the deceased and those institutionalised). Of these, 8,780 were successfully interviewed at Wave 2.

The response analysis was applied at the *individual-level* of analysis (although a number of household-level variables were included in the model). Since all cases were interviewed at Wave 1, there were a large number of variables available for analysing response at Wave 2.² The dependent variable was a response status indicator (1 indicating a successful interview at Wave 2 and 0 indicating non-

¹ If non-core sample members are to be analysed they should be analysed unweighted.

² The analysis of response was performed on data weighted by the Wave 1 weight (**W1WGT**) so that the subsequent adjustment of this weight at Wave 2 did not replicate the Wave 1 weight.

response). The analysis showed the following groups as being the most likely to have *not* been successfully interviewed at Wave 2:

- not interviewed at HSE,
- limiting longstanding illness at HSE,
- head of household in the lower supervisory & technical, semi-routine or other National Statistics Socio-Economic Classification (NS-SEC) at Wave 1,
- living in London at Wave 1,
- sampled from HSE 1999 (rather than 1998 or 2001),
- non-white ethnicity,
- renting or 'other' housing tenure compared to owner-occupiers at Wave 1,
- marital status of single (never married) or married (first and only marriage) at Wave 1,
- CSE/other or no educational qualifications compared to those with a degree or equivalent at Wave 1,
- were current smokers at Wave 1,
- females aged 85 years and over at Wave 1.

The predicted probability of responding at Wave 2 for the responding individuals was inverted to provide the non-response adjustment at Wave 2. The weight was trimmed/truncated at the 99th percentile to avoid extreme weights having an undue influence on the estimates. It was then re-scaled so that the non-response weight was standardised around an average of 1.

Finally, the non-response weight was then multiplied by the Wave 1 weight. This weight (**W2WGT**) was also then scaled to ensure an average of 1 (and that the weighted total sample size equalled the unweighted).

(Note that calibration weighting has *not* been performed on the Wave 2 weight. One reason is that possible external sources of information for calibration (e.g. mid-year household population estimates or 2001 Census totals 'aged-on' to the mid-point of Wave 2 fieldwork) would have required an adjustment to exclude immigrants. By definition, immigrants cannot have entered the ELSA 'population' between Waves 1 and 2 as only core members interviewed at Wave 1 were followed-up for interview at Wave 2. We were advised by the Demographic Methods Centre at the Office for National Statistics that such an adjustment would have inevitably introduced some errors into the weighting).

Standard Errors

Standard errors for survey estimates should account for the complex sample design and the weights. There are two aspects of the sample design that impact on standard errors: clustering and stratification. Clusters can be identified with variable **AHSECLS2** (this is currently not in the Wave 2 dataset, but is in the Wave 1 dataset). The stratification can be replicated using variable **ASTRATIF2**.

The ELSA Wave 1 Technical Report illustrates the difference between true standard errors (i.e. those that take the design into account) and uncorrected standard errors. Design factors for key estimates tend to fall within the range 1.0 - 1.7, where values above 1.2 are commonly described as indicating sizeable variance inflation.

SURVEY CONTENT

As at the previous wave, the Wave 2 main survey comprised a personal face-to-face interview and a self-completion questionnaire. Overall, the intention in Wave 2 was to collect data about the same topics as in Wave 1. There were, however, some

additions to the content of the interview to respond to new areas of enquiry. Some questions from Wave 1 were omitted as it was decided that they did not need to be asked at every wave. Furthermore, several elements of the questionnaire were amended to take account of responses given at the previous wave.

The ELSA program allowed flexibility in administering the interview. Respondents could be interviewed individually, or in houses with more than one eligible respondent, interviewed at the same time (in a single session) using concurrent interviewing techniques. In a concurrent session (see Survey Module Identifiers section below) the same block of questions was asked alternately of each person. Concurrent interviews tended to be quicker than two separate individual interview sessions, and were generally more convenient for respondents.

ELSA Wave 2 adopted the use of dependent interviewing in many of its modules. The technique is also called 'feeding forward' data. It is a technique that feeds responses individuals made at earlier interviews to either aid recall and/or improve consistency of responses across interviews. Earlier responses can either be provided to the individual before they respond to a question (proactive) or can be used after the individual has responded as a form of validation (reactive). Dependent interviewing was used to check some information collected during Wave 1 (or HSE), to determine changes in status, and to control routing within the questionnaire.

There were various modules each covering a different area of interest. The content and major routing of each module is described below. Although interviews tended to follow the same module order, some flexibility was given to the interviewer. For example, the timed walk could be administered at any time after the Health module, and it was possible for interviewers to skip the Income and Assets or Housing modules if it was more convenient to do them at another time.

Five of the modules (Cognitive Function, Expectations, Effort and Reward, Psychosocial Health and Final Questions) form the 'private modules' block. Wherever possible, these modules were administered with no other household members present. If two respondents were being interviewed concurrently, whilst the first respondent was being asked the private block, the second responding individual was asked to fill in the self-completion in a separate room. The two respondents then switched places.

The questionnaire modules are listed below in the order programmed in the CAPI interview. Most of the names of the variables (including derived variables) in each of the modules start with the same two-letter module prefix (shown in the brackets below). However, there are a small number of variable names that do not follow this convention.

Household Demographics ("DH")

The household demographics module was answered by one person on behalf of the household. It was used to collect basic demographic information about everyone living in the household. The composition of the household given at Wave 1 was proactively fed forward and the interviewer checked if all the people were still present in the household, and if their details were correct. The ELSA program determined the eligibility for the ELSA interview for each person in the household (see Sample Design section).

Note that information provided by the respondent to these questions is copied onto the data for other members of the household.

Individual Demographics (“DI”)

This module is at the start of the ELSA individual interview. Each respondent was asked details about their legal marital status, living children including adopted, foster and stepchildren, number of grandchildren and great-grandchildren, number of siblings and their own circumstances in childhood (if not answered at Wave 1).

Health (“HE”)

This module was administered to all respondents. It covered many different dimensions: self-reported general health: long-standing illness or disability; eyesight and hearing; specific diagnoses and symptoms; pain; difficulties with activities of daily living (ADLs) and instrumental activities of daily living (IADLs); and health behaviours. Respondents aged 60 and over were asked about falls and fractures. New questions were added about the quality of healthcare ELSA respondents have received for particular health conditions.

Social Participation (“SP”)

This module was administered to all respondents. It was much shorter than the Wave 1 SP module, and only asked about the use of public transport. The questions on caring for others were moved to the new 'Effort and Reward' module and the questions on cultural capital (e.g. how often the respondent eats outside the house or goes to the theatre) were asked in the self completion questionnaire instead.

Work and Pensions (“WP”)

This module was administered to all respondents. The aim of this module was to update or obtain information on the respondent's current work activities, if any, and update or obtain information about any new or current pensions that they may have.

Income and Assets (“IA”)

At the start of the interview couples were asked whether they kept their finances together or separate (**DHIASEP** – this variable is not in the data). If kept together, they were considered to be a single financial unit which required only one respondent for the IA module (the “financial respondent”). The couple decided who the respondent would be, although the interviewer was instructed to suggest to the couple that the person who answered the IA module at Wave 1 did so again at Wave 2. If their finances were kept separately, each person needed to answer the IA module separately and were treated as two separate financial units.

The type of financial unit the respondent was in is stored in a variable called **FUTYPE**. The variable which identifies who answered the questions in IA is called **IAPID**. Note that the interviewer can choose to answer IA at a later time. This added flexibility for respondents and interviewers and inevitably means that a small number of people did not answer IA at all. **IAPID** takes the value -1 when no one in the financial unit answered IA.

The information provided by the financial respondent is copied across to the other member of the financial unit (where this is appropriate). One important implication of this is that where a question refers to a person's “partner” or “spouse” (e.g. **IASPP** “how much does your husband/wife/partner receive from the state pension”), for individuals whose partners answered IA on their behalf, the data recorded in this variable will actually be referring to them (rather than their partner).

Details of the income that respondents received from a variety of sources over the last 12 months were collected including wages, state pensions, private pensions, other annuity income and state benefits. In addition, this module covered the amount

of financial and non-financial assets held, any income from these assets, regular transfers from non-household members and one-off payments in the last year.

Housing ("HO")

Only one eligible ELSA respondent in the household answered the housing module. Respondents decided themselves who the housing respondent should be, but again, the interviewer was instructed to suggest that the person who answered the HO module at Wave 1 answered this module again at Wave 2. Information provided by the HO respondent is copied across to other members of the household. The variable **HOPID** identifies who answered HO on behalf of the household.

This module collected or checked information about respondent's current housing situation, such as mortgages, value of property, rent etc, housing-related expenses, and ownership of cars. New questions were added on durable purchases (including amount spent) in the last 2 years, money spent in the last four weeks on clothes, leisure activities and gifts, and fuel expenditure.

Cognitive Function ("CF")

This module was the start of the 'private modules' block, i.e. those which were administered with no other household members present. The CF module was asked of all respondents (except proxies – see Proxies section) and measured different aspects of the respondent's cognitive function, including memory, speed, and mental flexibility. Questions to assess respondents' literacy skills replaced the numeracy questions used in Wave 1, although numeracy will be assessed again in future waves.

Expectations ("EX")

This module was asked privately of all respondents. It measured people's expectations in a number of dimensions, the level of certainty respondents felt about the future, financial decision-making within households and the time frame they thought about when making financial decisions. In Wave 2 respondents were also asked about the extent to which they expect to move out of their current home at some time in the future, and those aged 60 or over were asked about the chances that they would move into a nursing home in the next five years. Respondents were asked two open questions about the most positive and negative things about growing older (although this data has not been archived as the answers are text variables – see the Dropped Variables section below).

Effort and Reward ("ER")

This module, which was asked privately of all respondents, was new to Wave 2, although some of the questions in it were part of the SP module at Wave 1. It assesses respondents' motivations for doing voluntary work and caring for others and the relationship between effort and reward.

Psychosocial health ("PS")

This module was asked privately of all respondents. This module assesses symptoms of depression, based on the CES-D depression scale. At Wave 2, respondents who reported depression symptoms on this scale were asked about the quality of healthcare they have received for depression. There are some additional psychosocial questions in the self-completion questionnaire.

Final questions ("FQ")

This module was asked privately of all respondents. It checked information given by respondents at the previous wave such as demographic information, qualifications, and the details of their "stable address contact" (i.e. someone who could be

contacted if the respondent had moved at the next wave to find out their new address). In addition, respondents were asked to nominate someone who could do a proxy interview for them at the next wave if they were unable to be interviewed in person (see Proxies section below). If respondents had given their consent at Wave 2 to link their survey data to health and economic data from administrative sources they were given a written reminder of this consent. Respondents who had not consented to this at Wave 1 were asked if they would consent at this wave.

Walking speed (or Measurement – “MM”)

This module could be conducted at any point in the interview after the HE module. The test of walking speed was completed by respondents aged 60 and over when it was judged safe to do so. Respondents were asked to walk a distance of 8 feet (244 cm) at their usual walking pace. This walk was performed twice by respondents, and the interviewer recorded the time taken using a stopwatch (see protocol in interviewer instructions).

Self-completion (“SC”)

When respondents completed a full interview in a session alone, the self-completion questionnaire was usually left at the end to be returned by the respondent by post. However, when two respondents completed the interview in a concurrent session, the self-completion questionnaire was completed by one respondent while the other carried out the “private” modules of the personal interview (i.e. CF, EX, ER, PS, and FQ).

The self-completion questionnaire contained measures of well being and quality of life and questions about the respondent’s social circumstances, for example their social participation, social networks and relationships. New questions were added at Wave 2 on social isolation, life satisfaction and alcohol consumption.

PROXIES

If an eligible respondent was physically or cognitively impaired, or in hospital or temporary care for the whole of the fieldwork period, a proxy interview was permissible. Interviewers were asked to identify a proxy informant (i.e. a person who could answer the interview on behalf of the eligible respondent). The proxy informant was a responsible adult (aged 16 years or over) who knew enough about the respondent’s circumstances to be able to provide information about them. Where possible, a close family member such as a partner or son or daughter fulfilled this role.

The CAPI program guided the interviewer through the proxy interview automatically. However, only a subset of questions was asked during a proxy interview. The proxy interview contained the following modules (the asterisked modules were not asked for all respondents – see below):

HD*	Household grid
ID	Individual demographics
HE	Health
WP	Work and Pensions
IA*	Income and Assets
HO*	Housing
FQ	Final questions and consents

All proxy informants completed ID, HE, WP and FQ. Some specific questions for proxies were included in the Health module, HE.

The modules on household demographics and housing were done as part of the proxy interview only when no-one else in the household was eligible for interview.

In cases where no-one else in the same financial unit was eligible for interview, the proxy interview included the Income and Assets module. If two proxy interviews were needed for a couple, the Income and Assets module would only appear in one of the interviews (asking about both their finances). For couples comprising of one person who was interviewed by person and another who needed a proxy interview, the former would automatically be asked the Income and Assets module on behalf of the couple. The question about whether they keep finances together or separate would not be asked.

The questionnaire documentation shows whether each question was asked of proxy and/or non-proxy respondents:

* - If a variable has ONE asterisk next to it, this means that the question was asked of both proxy and non-proxy respondents

** - If a variable has TWO asterisks next to it, this means that the question was ONLY asked of proxy respondents

If a variable has NO asterisks next to it, this means that the question was NOT asked of proxy respondents

The variable **ASKPX1** identifies proxy interviews. 92 proxy interviews were for core members, 19 were for core partners, 6 were for new partners, and 8 were for younger partners.

USE OF RESPONSES FROM PREVIOUS INTERVIEWS IN THE CAPI (FEED FORWARD)

In the ELSA Wave 2 program, answers given at previous interviews (i.e. Wave 1 or HSE) were used in the routing and question wording (i.e. textfills). For example, if a respondent reported that they had an eye condition (e.g. cataract) at Wave 1, they were asked particular questions at Wave 2 to find out if they still had this condition (**HEOPC, HEOPW, HEOPN, HEOPS**).

Problem with feed forward data - HEDIAC

For the first part of the Wave 2 fieldwork there was a problem with the variable **HEDIAC**, which asked respondents who had reported a cardiovascular condition at Wave 1 whether they still have it at Wave 2. This question was not asked of any respondents in the first part of fieldwork. Unfortunately, this also had a knock on effect on many other follow up questions about cardiovascular conditions (see Notes about particular variables in the data on Page 19 of the user guide).

The variable **HEFFLAG** shows which respondents reported one or more cardiovascular conditions at Wave 1 who were not asked HEDIAC due to the feed forward problem (N=2532).

Problem with feed forward data - wrong information fed forward

Furthermore, in a small number of cases (N=33), due to a problem with the CAPI program, the wrong set of previous responses was fed forward for various questions. Therefore, these respondents were asked a few questions that they should not have been asked, were not asked some questions that they should have been asked, or were asked questions with the wrong wording (i.e. textfills). Only a small number of questions were affected by this problem. This is partly because the (wrong) responses that were fed forward usually came from another member of the household, who had often given the same answers as this person at Wave 1.

The respondents who had the wrong information fed forward are identified by the variable **FFPROB**. These respondents have also had their answers to particular questions recoded to show that their responses to these questions were affected by this problem. The two special response categories that have been used are:

- 3 'Did not respond at w1 but were not asked'
- 4 'Did respond at w1 but were asked'

DATASET INFORMATION

The core dataset has been archived for ELSA Wave 2. This lists cases at an individual level.

Data for ELSA respondents collected during HSE (referred to as "Wave 0" data) and ELSA Wave 1 can be used for longitudinal analysis, all of which is available from the data archive.

The ELSA Wave 2 core dataset contains data in the following order:

- Key variables not in the questionnaire (e.g. serial number, outcome codes)
- Variables in the questionnaire (in the order they appear in the CAPI interview). A small number of additional variables that are associated with particular questionnaire variables are located alongside these variables in the data. These are annotated in the questionnaire.
- Other variables not in the questionnaire (including administrative variables, other derived variables and variables fed forward from Wave 1).

See Appendix for the full list of variables not included in the questionnaire.

The main group of respondents for analysis is the core members. Data on partners (i.e. core, younger and new partners) can be used as characteristics of the core members (i.e. to provide supplementary information), but these partners should not be analysed as individuals in their own right. The ineligible partners are unrepresentative, and any analysis using them would need to be unweighted.

SERIAL NUMBERING

Constant serial numbers

All the ELSA data files deposited in the archive will contain a unique individual analytical serial number (**IDAUNIQ**) to enable users to link the different files. Each respondent will have a unique value for **IDAUNIQ** which will remain constant.

Wave-specific household serial number

The five digit household serial number (**IDAHHW2**) was randomly generated for the archived dataset, and does not relate to the serial number used during interviewing. The dataset for each wave of ELSA (including "Wave 0", i.e. HSE) contains a different set of household serial numbers (**IDAHHW0**, **IDAHHW1** etc.). It is necessary to have a different household serial number for each wave as respondents can change households between waves. An index file will shortly be provided which will enable data users to link the household serial numbers in order to compare data for each respondent and household at different waves.

Wave-specific individual serial number

IDAINDW2 is the individual serial number. This serial number is independent of the household serial number and the serial number used during interviewing.

Person number

Each person within the household was given a number, starting from 1, at the time of the HSE interview (**PERID**). The numbering was continued for new people that entered the household after the HSE interview. This number is used for several variables. For example it is used to indicate which one of the other people in the household is the respondent's partner (see variable **CPID**), or to identify the person in the household that responded to particular questions such as the Income and Assets module (see variable **IAPID**).

INTERVIEW OUTCOME

Individual and household level interview outcomes are given in **W2INDOUT** and **W2HHOUT** respectively.

The individual outcome is a two-digit code. The first digit indicates whether the interview was full (1) or partial (2). The second digit indicates whether the interview was in person (1) or by proxy (3).

The household outcome is a three-digit code. The first digit indicates whether all eligible residents were interviewed (1) or some interviewed (2).

SURVEY MODULE IDENTIFIERS

Some modules in the interview were answered by all interviewed individuals whereas others were not. Each of the modules that could be answered by one individual on behalf of others has a variable that identifies the person who answered the module:

- Household Demographics = **HHRESP**
- Income and Assets = **IAPID**
- Housing = **HOPID**

Where modules were answered by one person on behalf of others, information was copied directly to the other member(s) of the household or financial unit.

The variable **CONCUR** indicates whether or not the respondent was interviewed alone (**CONCUR=2**) or concurrently with another respondent in the household (**CONCUR=1**). The person number of the person interviewed first in the concurrent session is provided in **ORDCON**.

SAMPLE TYPE

FINSTAT indicates the sample type of the individual:

- C1CM = Core sample member
- C1CP = Core partner
- C1YP = Younger partner (who responded at Wave 1)
- C1YP_unprod = Younger partner (who did not respond at Wave 1)
- C1NP1 = New partner (found at Wave 1 who responded at Wave 1)
- C1NP1_unprod = New partner (found at Wave 1 who did not respond at Wave 1)
- C1NP2 = New partner (found at Wave 2)

AGE

DHDOBYR (referred to as **DHDOBY** in the questionnaire) and **DHAGER** provide the date of birth and age of respondents recorded in the household grid. Note that a given respondent may not have provided this information themselves (as anyone in the household can complete the household grid). All respondents aged 90 or over have been given the following values for these variables for confidentiality reasons: **DHDOBYR=-7**; **DHAGER=99**.

DIAGR (referred to as **DIAG** in the questionnaire) is the age given in the individual interview. Again, all respondents aged 90 or over have been given a value of 99 for this variable for confidentiality reasons. **DIAGR=99**.

INDOBYR is derived from the date of birth variables from the household grid (**DHDOB**, not archived) and individual interview (**DIDBN**, not archived). If the variable from the individual session was answered (**DIDBNY**) then this value was used, otherwise the value from the household grid (**DHDOB**) was used in the derivation. **INDOBYR** provides the year of date of birth only; the day and month of birth have been dropped from the dataset to retain confidentiality.

INDAGER is computed from date of birth (**INDOB** – dropped from dataset) and date of interview (**INTDAT** – dropped from dataset). All respondents over age 90 have been classified as 99 years old for confidentiality reasons.

For age analysis, **INDOBYR** and **INDAGER** are the suggested variables to use.

GENDER

DISEX was the sex given in the respondent's individual questionnaire session. **DHSEX** was given in the household demographics module. **INDSEX** is derived from these two variables. If the variable from the individual session was answered (**DISEX**) then this value was used, otherwise the value from the household grid (**DHSEX**) was used in the derivation.

EDUCATION

If a respondent's qualifications had been collected during the HSE or ELSA Wave 1 interviews, only additional qualifications gained since the last interview were recorded at ELSA Wave 2. Those not interviewed previously, were asked for all qualifications. The qualifications collected during the ELSA Wave 2 interview are in **FQQUZM1** to **FQQUZM3**. If a respondent's highest qualification is needed, then a combination of Wave 1 and Wave 2 answers will likely be needed.

FINANCIAL VARIABLES

ELSA contains very detailed information on all aspects of finances. Summary variables have been derived and are available in a separate dataset ("ELSA Wave 2 Financial Derived Variables"). To derive these variables, some imputation was carried out and the methods used are described in the document "Financial Derived Variables User Guide and Imputation Procedures". Key variables contained in this supplementary dataset are:

Totinc_bu	benefit unit	total income
Empinc_bu	benefit unit	employment income
Seinc_bu	benefit unit	self-employment income
Ppinc_bu	benefit unit	private pension income
Spinc_bu	benefit unit	state pension income
Beninc_bu	benefit unit	benefit income
Assinc_bu	benefit unit	asset income
Othinc_bu	benefit unit	other income
Eqtotinc_bu	benefit unit	equivalised total income
Eqempinc_bu	benefit unit	equivalised employment income
Eqseinc_bu	benefit unit	equivalised self-employment income
Eqppinc_bu	benefit unit	equivalised private pension income
Eqspinc_bu	benefit unit	equivalised state pension income
Eqbeninc_bu	benefit unit	equivalised benefit income

Eqassinc_bu	benefit unit	equivalised asset income
Eqothinc_bu	benefit unit	equivalised other income
Savings_bu	benefit unit	total savings
Invests_bu	benefit unit	total investments
Grossfw_bu	benefit unit	total gross financial wealth
Debt_bu	benefit unit	total debt
Netfw_bu	benefit unit	total net financial wealth
Netpw_bu	benefit unit	total net physical wealth
Grosstotnhw_bu	benefit unit	total gross non-housing wealth
Nettotnhw_bu	benefit unit	total net non-housing wealth
Ghwealth_bu	benefit unit	total gross housing wealth
Nhwealth_bu	benefit unit	total net housing wealth
Nettotw_bu	benefit unit	total net (non-pension) wealth

DROPPED VARIABLES

All variables in the questionnaire documentation with a @ symbol next to their name have been deleted from the archived dataset.

The following types of variables have been deleted in order to reduce the potential to identify individuals and for other reasons (specified below):

1. Those containing text
2. Those which contained a personal identifier (e.g. name/address)
3. Those considered to be disclosive, such as:
 - Detailed ethnicity
 - Specific country of birth
 - Full interview date
 - Full date of birth
 - Council tax payments (Different councils charge different amounts and therefore the amount may reveal the area the respondent lives in)
 - Water and sewerage charges (These vary in different areas and therefore the amount may reveal the area the respondent lives in)
4. Timing variables, which give the time at specific points in the interview (used for administration purposes)
5. Variables that only contain missing values – excluded because not useful.

There are no geographical variables in the archived dataset. A separate geographical dataset with broad identifiers is available under secure arrangements. Please contact the ELSA data manager for more details (see Page 17 of the User Guide).

CODING AND EDITING

Additional coding and editing tasks were performed after the interviews were conducted. The ELSA Wave 2 Code Book and Edit Instructions document (part of the main User Guide) provides details of the tasks that were conducted.

CODING

The coding of responses was mostly dealt with by the CAPI questionnaire through the use of fully closed questions. There were, however, a number of questions (e.g. **WPHLWHO**) where a code frame was not used in order to capture all responses to a

particular question verbatim. The responses to these ‘open’ questions were coded into separate variables after the interview was conducted (e.g. open responses to **WPHLWHO** were coded into **WPHLWZ**). The coded responses to open questions are in the core dataset, but as mentioned above the original text responses have been removed to reduce the potential to identify individuals.

Other questions in the CAPI questionnaire had code frames which included an ‘other’ option (e.g. **DIKLIV**). In these ‘other-specify’ questions, interviewers could use this option if the respondent’s answer did not fit any of the codes or if they were not confident of coding into the prescribed codes. In these cases, the interviewer recorded the full ‘other’ answer at a follow up question (e.g. ‘other’ responses relating to **DIKLIV** were recorded in **DIKLO**).

If these other-specify questions were ‘single coded’, i.e. when only 1 option could be chosen (e.g. **DIKLIV**) then the text answers were coded and incorporated into the original coded variable (i.e. **DIKLO**).

For all multi-coded³ variables that were coded, there are two sets of variables. The first are the original variables that contain the answers recorded by the interviewer (e.g. **SPTRAB1** to **SPTRAB7**). The second set of variables contain the original coding plus the codes assigned to “other answers” (e.g. **SPTRM01** to **SPTRM07**). Note that the suffix ‘m’ added to these variable names means that they contain ‘merged’ original and coded answers. This naming convention was followed consistently so the final merged variables can be identified by name. The merged variables should be used instead of the original variables.

The original variables provide multi-coded answers in the order in which they were mentioned by the respondent. Please note that in some cases the order of the respondent’s answers in the merged variables is not the same as in the original variables.

EDITING

As with the coding, most of the editing for the ELSA Wave 1 questionnaire was carried out by the interviewers in the field. However, there were some additional checks that related to inconsistencies in the data which were carried out after the interview.

MISSING VALUES

For most questions there are the following missing values:

- 1 Not applicable
- 8 Don’t know
- 9 Refusal

For some questions, a response of ‘don’t know’ or ‘refusal’ was not permitted. This is indicated in the questionnaire.

For various reasons, some respondents did not complete the questionnaire. Where possible, for these ‘partial’ interviews, the questions that were not asked were coded as –1 (not applicable). This will be checked and refined in future versions of the dataset.

³ A variable is multi-coded if more than one response can be given e.g. favourite activities are reading and going to the cinema.

CONTACT DETAILS

ELSA Data Manager:

Susan Nunn

E-mail: s.nunn@natcen.ac.uk

Telephone: 020 7250 1866

NOTES ABOUT PARTICULAR VARIABLES IN THE DATA

Please note that you should also look at the questionnaire to find out the exact wording of the question and response options, who was eligible for each question and other information.

Variable name	Variable label	Comments
HEATA	Main symptom making walking 1/4 mile difficult	This question is asked if respondent reported 2 or more symptoms at HEATT. To obtain the main symptom for all respondents, combine HEATA with symptom given at HEATT01 by those who only had one symptom.
HEBALA1, HEBALA2	Balance 1: whether did treatment after recommended by doctor/nurse (walking symp)	<p>These questions, about the quality of care respondents have received for balance problems, occur twice in the interview – (i) after questions on walking 1/4 mile (HEATT) and (ii) after direct questions on balance (HEBAL). In Wave 2, the questions in the two subsets were in a different order and had different routing and therefore cannot be combined.</p> <p>HEBALA1 asks only if known to have done exercises or had physiotherapy and asks whether the action was taken after recommendation by a doctor or nurse. HEBALA2 asks if a doctor or nurse recommended the action, before asking whether the respondent followed the recommendation</p>
HEBALB1, HEBALB2	Balance: whether stick or zimmer recommended by doctor/nurse (walking symptom)	As above, these questions occur twice in the interview – (i) after questions on walking 1/4 mile (HEATT) and (ii) after direct questions on balance (HEBAL). In Wave 2, the questions in the two subsets were in a different order and had different routing and therefore cannot be combined.
HEBALC1, HEBALC2,	Balance: whether stick or zimmer recommended by doctor/nurse (walking symptom)	As above.
HEAID1-5 HEAID9-12 HEAID17-22	Walking aids used	<p>The question about what walking aids the respondent uses is asked at three points in the interview:</p> <ul style="list-style-type: none"> (i) after the questions on walking 1/4 mile (HEATT) – answers in HEAID1-5 (ii) after questions on balance (HEBAL) – answers in HEAID9-12 (iii) after questions on activities of daily living (HEADLA01) – answers in HEAID17-22 <p>Each respondent was only asked the HEAID question once.</p>
HEOPC	Whether confirms diagnosed eye condition recorded in Wave 1	This question asked respondents about the eye conditions they said they had at Wave 1 at this question. If the respondent disputes having one or more of these conditions at this question, the disputed conditions are recorded at HEOPW1-HEOPW2.
HEOPS	Whether still has eye condition at Wave 2	This question was asked of respondents who reported having an eye condition at Wave 1 which they did not dispute at HEOPW1-HEOPW2 at Wave 2 (see questionnaire for exact routing). However, this

		is provided they gave a response at HEOPC.
HEOPT1- HEOPT2	Diagnosed eye condition newly reported at W2	<p>For all the eye conditions, except cataracts, respondents were not asked about conditions they reported at Wave 1. If respondents reported having cataracts at Wave 1, they were still asked about cataracts again at this question to find out if they had a new diagnosis since their last interview. However, the reporting of additional cataracts may be incomplete.</p> <p>Respondents who answered this question at Wave 1 were asked if they had been diagnosed within the last two years, whereas respondents who did not answer this at Wave 1 were asked if they had ever been diagnosed.</p>
HEBPCHK	Blood pressure: whether checked by doctor or nurse in past year	There was a routing error so this was not asked of the right age group throughout the period.
HEDIAW1- HEDIAS9	(Questions about CVD conditions reported at Wave 1)	These variables check the CVD condition(s) which the respondent reported at Wave 1. These variables need to be used in conjunction with HEDIA01-09 (named HEDIAA01-09 in the questionnaire) in order to obtain the number of people who have ever been diagnosed with a CVD condition. Please note that there was a routing error for HEDIAC at the beginning of the fieldwork which meant that some people were not asked about the CVD condition(s) they reported at Wave 1. (See HEFFLAG and Section on Use of responses from previous interviews in the CAPI (feed forward)). Therefore, users should merge in Wave 1 data to obtain the full number of those ever-diagnosed.
HEDIAC1- HEDIAC9	Whether confirms [CVD condition] recorded in Wave 1	Respondents were asked to confirm each of the CVD conditions they reported having at HEDIAA at Wave 1. There is a separate variable for their responses about each of the CVD conditions. So, for example, responses about whether the respondent confirmed the first condition (high blood pressure) recorded in Wave 1 is recorded in HEDIAC1. (See HEFFLAG and Section on Use of responses from previous interviews in the CAPI (feed forward)).
HEDIAN1- HEDIAN9	Reason disputed [CVD condition] diagnosis from Wave 1	If the respondent disputes a CVD condition they reported at Wave 1 they are asked the reason for this dispute at HEDIAN. As with HEDIAC, there is a separate variable for each condition. (See Section on Use of responses from previous interviews in the CAPI (feed forward)).
HEDIAS1- HEDIAS9	Whether still has [CVD condition] at Wave 2	If the respondent confirmed the CVD condition they reported at Wave 1 at HEDIAC, they were asked if they still had this condition at HEDIAS. As with HEDIAC, there is a separate variable for each condition. However, there is no variable HEDIAS3 or HEDIAS8 as respondents were not asked if they still had a heart attack or stroke as this would not make sense. Note that answers to this question refer to the respondent's perceptions and do not necessarily reflect any clinical judgement that they are cured. (See Section on Use of responses from previous interviews in the CAPI (feed forward)).

HEFFLAG	Reported 1+ CVD condition at W1 but not asked HEDIAC due to feed forward problem	At the beginning of fieldwork there was a routing error for HEDIAC in the CAPI interviewing program which meant that some people were not asked about the CVD condition(s) they reported at Wave 1. HEFFLAG indicates whether or not a respondent had a condition which should have been feed forward but was not (See Section on Use of responses from previous interviews in the CAPI (feed forward)).
HEDIA01- HEDIA09 (named HEDIAA01- HEDIAA09 in the questionnaire)	Diagnosed cardiovascular condition newly reported at W2	For all the CVD conditions, except heart attacks and strokes, respondents were not asked about conditions they reported at Wave 1 at this question. If respondents reported having a heart attack or stroke at Wave 1, they were given the opportunity to report a fresh diagnosis at Wave 2. However, users should also check HENMMI, HEYRA, and HEYRB for recent strokes or heart attacks. Respondents who answered this question at Wave 1 were asked if they had been diagnosed within the last two years, whereas respondents who did not answer this at Wave 1 were asked if they had ever been diagnosed.
HEHBP (named HEHIBPB in the questionnaire) HEHBPB1 (named HEHIBPB1 in the questionnaire)	Whether doctor or nurse suggested medication to lower blood pressure	This question is asked in two places, but only once of each respondent: (i) In the follow-up questions to diagnosed high blood pressure (HEDIAA01-9 in questionnaire, HEDIA01-9 in the data) – recorded in HEHBP) (ii) After the questions about strokes (HEAGE) - recorded in HEHBPB1
HEAGA HEAGB HEAGC HEAGD HEAGE	Age first told had [CVD condition]	At Wave 2 these questions were asked only of the following: (i) respondents who were not interviewed at Wave 1 and said they had the condition at Wave 2 (HEDIA01-9) (ii) respondents who had the condition at Wave 1 and confirmed this at Wave 2 but did not respond to the question about at what age they were diagnosed with it at Wave 1 See below as well.
HEAGAR & HEAGARY, HEAGCR & HEAGCRY, HEAGDR & HEAGDRY	Month and year told had [CVD condition] in last 2 years	At Wave 2 these questions were asked only if respondents did not have the condition at Wave 1 but did have it Wave 2. Use these variables in conjunction with the age of diagnoses reported at Wave 1 (BHEAGA-BHEAGE) and HEAGA-HEAGE to obtain the age at diagnosis for everyone reporting diagnosis up to Wave 2.
HEAGBR & HEAGBRY, HEAGER & HEAGERY,	Month and year told had [CVD condition] in last 2 years – heart attack and stroke	These questions were asked of everyone who responded at HEDIAA at Wave 1 and reported a new condition at this question at Wave 2.
HEHRTA- HEHRTF	(Questions about quality of care received for angina or heart attacks)	These questions were asked of two sets of respondents: (i) those who reported having being diagnosed with angina or heart attack (ii) those who did not report diagnoses of these

		<p>or stroke or diabetes but showed angina symptoms in the Rose angina questionnaire and said that they had talked to a doctor or nurse about them (HEANL).</p> <p>Due to a problem with the routing for HEDIAC, this information is not available for some people whose diagnosis was reported at Wave 1 (See HEFFLAG and Section on Use of responses from previous interviews in the CAPI (feed forward)). Although, it is available for those who reported at HEDIAC at Wave 2.</p>
HECHOL- HECHOLC	(Questions about quality of care received for high cholesterol)	<p>HECHOL was asked of two sets of respondents:</p> <ul style="list-style-type: none"> (i) those who reported having a diagnosis of high cholesterol AND angina, heart attack, diabetes (ii) those who did not report a diagnosis of high cholesterol but did report pain in their sternum, arm or chest at HEANI <p>Due to a problem with the routing for HEDIAC, this information is not available for some people whose angina or heart attack was reported at Wave 1 (See HEFFLAG and Section on Use of responses from previous interviews in the CAPI (feed forward)).</p>
HEDIBW1- HEDIDS9	(Questions about chronic conditions reported at Wave 1)	These variables check the chronic condition(s) which the respondent reported at Wave 1. These variables need to be used in conjunction with HEDIB01-09 (named HEDIAB01-09 in the questionnaire) in order to obtain the number of people who have ever been diagnosed with one of the chronic conditions.
HEDIAD1- HEDIAD9	Whether confirms [chronic condition] recorded in Wave 1	Respondents were asked to confirm each of the chronic conditions they reported having at HEDIAB at Wave 1. There is a separate variable for their responses about each of the chronic conditions. So, for example, responses about whether the respondent confirmed the first condition (chronic lung disease) recorded in Wave 1 is recorded in HEDIAD1.
HEDIAM1- HEDIAM9	Reason disputed [chronic condition] diagnosis from Wave 1	If the respondent disputes a chronic condition they reported at Wave 1 they are asked the reason for this dispute at HEDIAM. As with HEDIAD, there is a separate variable for each condition.
HEDIDS1- HEDIDS9	Whether still has [chronic condition] at Wave 2	If the respondent confirmed the chronic condition they reported at Wave 1 at HEDIAD, they were asked if they still had this condition at HEDIDS. As with HEDIAD, there is a separate variable for each condition.
HEDIB01- HEDIB09 (named HEDIAB1- HEDIAB9 in the questionnaire)	Diagnosed non-cardiovascular condition newly reported at W2	<p>Conditions already reported at Wave 1 are not normally re-reported at this question, which is intended to ask about additional diagnoses.</p> <p>Respondents who answered this question at Wave 1 were asked if they had been diagnosed within the last two years, whereas respondents who did not answer this at Wave 1 were asked if they had ever been diagnosed.</p>
HEARTA- HEARTE	(Questions about quality of care received for	Due to a problem with the routing, these questions were only asked for incident cases (i.e. if reported diagnosis at Wave 2 rather than Wave 1).

	osteoarthritis)	
HEPSY1-6	Type of psychiatric problem	This question was asked of people who reported having emotional, nervous or psychiatric problems at HEDIB01-9 at Wave 2. HEPSY1-6 may have been answered by people who reported having a psychiatric condition at Wave 1, as these respondents were allowed to report additional psychiatric problems at HEDIB01-9 at Wave 2.
HEPSYA to HEPSYE		These questions were asked of two subsets of people: (i) those who reported having diagnosed depression at HEPSY (HE module) (ii) those who reported symptoms of depression which they had discussed with a doctor (PS module – PPSYA-PSPSYE). Respondent were only asked this set of questions once although they may fulfil both criteria.
Benefit receipt questions (e.g. IAA)	Benefit receipt questions	Care should be taken when using the benefit receipt questions between waves 1 and 2 because variables that have the same name do not necessarily relate to the same benefit because of changes to the benefit system between years.
PSFEEL-PSPSYE	(Questions about quality of care received for depression)	Respondents were asked these questions if they reported five or more depressive symptoms from the eight-item CES-D scale (psceda-pscedh). This was done on a listwise deletion basis, i.e. people who refused to answer or reported 'don't know' on any of the CES-D questions were excluded from being asked the quality of care questions.
SCACTA	How often respondent goes to the cinema	In the first wave of ELSA this variable was included in the CAPI (SPCin)
SCACTB	How often respondent eats out of the house	In the first wave of ELSA this variable was included in the CAPI (SPREst)
SCACTC	How often respondent goes to an art gallery or museum	In the first wave of ELSA this variable was included in the CAPI (SPMus)
SCACTD	How often respondent goes to the theatre, a concert or the opera	In the first wave of ELSA this variable was included in the CAPI (SPTea)
SCMOREA	Would respondent like to go to the cinema more often	In the first wave of ELSA this variable was included in the CAPI (SPCinB)
SCMOREB	Whether would like to eat out of the house more often	In the first wave of ELSA this variable was included in the CAPI (SPREstB)
SCMOREC	Whether would like to go to art galleries or museums more often	In the first wave of ELSA this variable was included in the CAPI (SPMusB)
SCMORED	Whether would like to go to the theatre, concerts or the opera more often	In the first wave of ELSA this variable was included in the CAPI (SPTeaB)

APPENDIX: Variables not included in the questionnaire documentation

Variable name	Variable label
BEGINNING OF DATASET	
IDAUNIQ	Unique individual serial number
IDAINDW2	W2 Individual analytical serial number
IDAHHW2	W2 household analytical serial number
PERID	Person number in the household
SAMPSTA	Type of sample member (i.e. core member, core partner, younger partner or new partner)
FINSTAT	Final status of respondent after Wave 2 interview
W2INDOUT	Individual outcome code
W2HHOUT	Household outcome code
COUPLE	Relationship status (derived from household grid)
CPID	Person number of partner
CHINHH1	Whether respondent has a child in the household
CHOUTH1	Whether respondent has a child outside the household
GCINHH1	Whether respondent has a grandchild in the household
MAINHH1	Whether the mother of the respondent is in the household
PAINHH1	Whether the father of the respondent is in the household
END OF DATASET	
CONCUR	Whether interview was concurrent or not
ORDCON	Person number of first person interviewed concurrently
WHOSO1	Anyone other than interviewer and respondent in the room (1st other person)
WHOSO2	Anyone other than interviewer and respondent in the room (2nd other person)
WHOSO3	Anyone other than interviewer and respondent in the room (3rd other person)
HHTOT	Number of people in household / computed
ASKPX1	Whether interviewed by proxy
FUTYPE	Financial unit type
INTDATM	Month of Household Interview
INTDATY	Year of Household Interview
IINTDTM	Month of Individual Interview
IINTDTY	Year of Individual Interview
INDSEX	Definitive sex variable. Priority: Disex, Dhsex
INDOBYR	Definitive year of birth collapsed at 90 plus. Priority: Didbn, Dhdob
INDAGER	Definitive age variable collapsed at 90 plus. Priority: Diag, Dhage
W1INDOUT	Wave 1 individual outcome code
W1HHOUT	Wave 1 household outcome code
AHSECLS2	Data Archive HSE clustering variable
HSEINT	Whether interviewed at HSE
HSEYR	HSE source survey year
QVER	Questionnaire version
FFPROB	Whether had feed forward data error (see user guide)
ASTRATIF2	HSE stratification variable (first level stratification)
ANSSEC	FROM HSE: NS-SEC - long version
AXSC2000	FROM HSE: SOC2000 (without dots)
BNSSEC	From Wave 1: NS-SEC
BXSC2000	From Wave 1: SOC2000 (without dots)
BHEAID1	Walking aids used reported in Wave 1 (1st mention)
BHEAID2	Walking aids used reported in Wave 1 (2nd mention)
BHEAID3	Walking aids used reported in Wave 1 (3rd mention)
BHEAID4	Walking aids used reported in Wave 1 (4th mention)
BHEAID5	Walking aids used reported in Wave 1 (5th mention)

BHEOPT1	Diagnosed eye condition reported at Wave 1 (1st mention)
BHEOPT2	Diagnosed eye condition reported at Wave 1 (2nd mention)
BHEOPT3	Diagnosed eye condition reported at Wave 1 (3rd mention)
BHEOPT4	Diagnosed eye condition reported at Wave 1 (4th mention)
BEHDIA01	Diagnosed cardiovascular condition reported at Wave 1 (1st mention)
BEHDIA02	Diagnosed cardiovascular condition reported at Wave 1 (2nd mention)
BEHDIA03	Diagnosed cardiovascular condition reported at Wave 1 (3rd mention)
BEHDIA04	Diagnosed cardiovascular condition reported at Wave 1 (4th mention)
BEHDIA05	Diagnosed cardiovascular condition reported at Wave 1 (5th mention)
BEHDIA06	Diagnosed cardiovascular condition reported at Wave 1 (6th mention)
BEHDIA07	Diagnosed cardiovascular condition reported at Wave 1 (7th mention)
BHEAGA	Age first told had angina reported at Wave 1
BHEAGB	Age first told had heart attack reported at Wave 1
BHEAGC	Age first told had congestive heart failure reported at Wave 1
BHEAGD	Age first told had diabetes or high blood sugar reported at Wave 1
BHEAGE	Age first told had stroke reported at Wave 1
BHEART1	Type of arthritis reported at Wave 1 (1st mention)
BHEART2	Type of arthritis reported at Wave 1 (2nd mention)
BHEART3	Type of arthritis reported at Wave 1 (3rd mention)
BHEAGF	Age first told had arthritis reported at Wave 1
BHEAGG	Age first told had cancer reported at Wave 1
BHECANA	Site that tumour/cancer started reported at Wave 1
BHEPRK	Age told had Parkinson's Disease reported at Wave 1
BHEAGH	Age told had psychiatric problems reported at Wave 1
BHEAGI	Age told had Alzheimer's Disease reported at Wave 1
BHEAGJ	Age told had dementia reported at Wave 1
BHESMK	Whether reported ever smoked cigarettes at Wave 1
BHESKA	Whether smokes cigarettes at all nowadays reported at Wave 1